軌道補正について（2）

M. Kikuchi

November 9, 2011

1　拘束条件付き線型方程式

1.1　問題の背景

ある拘束条件の下に方程式を解く必要に迫られることがしばしばある。この小論では方程式、拘束条件いずれも線型の場合を扱う。軌道補正を例にとれば、(a_{ij}) をステアリング-BPM 系の応答行列とす

れば、

$$
\begin{aligned}
\sum_{j=1}^{m} a_{ij}\theta_j &= y_i \ , \quad (i = 1, \ldots, n) \ , \qquad (1) \\
a_{ij} &= \sqrt{\beta_i \beta_j} \frac{\cos(\pi\nu - |\psi_i - \psi_j|)}{2 \sin \pi\nu}
\end{aligned}
$$

となる。θ はステアリングのキック角、y は BPM の読み値である。通常 BPM の数の方がステアリングの数よりも多いので、式 (1) の左辺と右辺の差の二乗和が最小になるようなキック角を求める。一方拘束条件は、例えば次のようなものが考えられる。

- エネルギー一定の条件：キック角の総和が零

$$\theta_1 + \cdots + \theta_m = 0$$

- ある点の軌道条件：

$$\sum_{j=1}^{m} a_{kj}\theta_j = y_k$$

- 周長一定の条件：(η は分散関数)

$$\sum_{j=1}^{m} \eta_j\theta_j = 0$$

- チューン一定の条件：(k' は六極磁石の強さ、N_s は六極磁石の個数)

$$\Delta\nu = \frac{1}{4\pi}\sum_i \beta_i k'_i \Delta x_i$$

$$\sum_{j=1}^{m}(\sum_{i=1}^{N_s} \beta_i k'_i a_{ij})\theta_j = 0$$

$$\Delta x_i = \sum_j a_{ij}\theta_j$$

これらの拘束条件を厳密に満たしつつ、式 (1) の左辺と右辺の差の二乗和が最小になるようなキック角を求めることが問題となる。したがってこの種の問題は一般に次のように定式化されるであろう。

1.2　問題

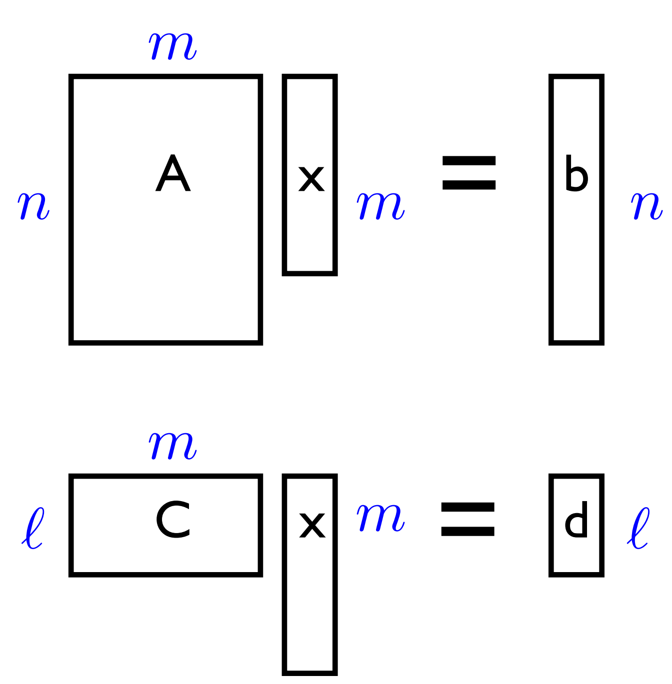
$$Ax = b \tag{2}$$

を拘束条件

$$Cx = d \tag{3}$$

の下に解くこと。ただし、A, x, b, C, d はそれぞれ次の型の matrix あるいは縦 vector である：

$$A(n, m),\ x(m),\ b(n),\ C(\ell, m),\ d(\ell) \tag{4}$$



2 Lagrange 未定乗数法

スカラー F を次の式で定義する：

$$F = \frac{1}{2}\|Ax - b\|^2 + \lambda^t (Cx - d) \tag{5}$$

λ は未定乗数（ここでは ℓ 次元縦 vector）である。拘束条件式 (3) の下に、F が最小になるような x が解である。x が F の停留値を与えることから、

$$\frac{\partial F}{\partial x} = 0\ . \tag{6}$$

式 (6) と式 (3) を同時に満足する x が解となる。具体的には式 (6) と式 (3) を x 及び λ について解くことになるが、こうして決められた λ は答えには直接影響しない。式 (6) を具体的に書くと、

$$\frac{\partial F}{\partial x} = A^t A x - A^t b + C^t \lambda = 0 \tag{7}$$

となる。式 (7)(3) から、方程式

$$My \equiv \begin{pmatrix} A^tA & C^t \\ C & 0 \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} A^tb \\ d \end{pmatrix} \equiv e \tag{8}$$

を得る。行列 M は $(m+\ell, m+\ell)$ の対称正方行列である。
$rank(M) = m+\ell$ のとき、式 (8) は一意的に解ける。
$rank(M) < m+\ell$ のとき、e が M の像に含まれれば解は存在する。
このとき $kerM$ だけ解には不定性が残るが、M の特異値分解によれば $\|y\|$ が最小となる解を得ることができる。

2.1　例

問題

$$\begin{cases} x_1 - x_2 + x_3 = 1 \\ -x_1 + x_2 - x_3 = 0 \end{cases}$$

を拘束条件

$$x_1 + x_2 + x_3 = 0$$

の下に解け。

答　Lagrange 未定乗数法で解いてみる。

$$M = \begin{pmatrix} 2 & -2 & 2 & 1 \\ -2 & 2 & -2 & 1 \\ 2 & -2 & 2 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} , \quad e = \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix} .$$

$$M = UWV^t ,$$

$$U = \begin{pmatrix} -0.58535 & 0.317283 & -0.238111 & 0.707107 \\ 0.551627 & 0.541713 & -0.634235 & 0 \\ -0.58535 & 0.317283 & -0.238111 & -0.707107 \\ -0.102169 & -0.710781 & -0.695954 & 0 \end{pmatrix} ,$$

$$W = \begin{pmatrix} 6.05932 & 0 & 0 & 0 \\ 0 & 1.65491 & 0 & 0 \\ 0 & 0 & 1.59559 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$V = \begin{pmatrix} -0.58535 & 0.551627 & -0.58535 & -0.102169 \\ -0.317283 & -0.541713 & -0.317283 & 0.710781 \\ -0.238111 & -0.634235 & -0.238111 & -0.695954 \\ 0.707107 & 0 & -0.707107 & 0 \end{pmatrix}.$$

$$M \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \lambda \end{pmatrix} = e$$

を解いて、

$$x_1 = 0.125,\ x_2 = -0.25,\ x_3 = 0.125,\ \lambda = 0$$

を得る。このとき $\|Ax-b\|=0.707107$ となる。

3 Schmidt 直交化法

式 (3) の一つの解を x_0 とすると、任意の解は $x=x_0+x'$ と書ける。x' は C の核に属するベクトルである。

$$x = x_0 + x' \;, \tag{9}$$

$$Cx_0 = d \;, \tag{10}$$

$$Cx' = 0 \;. \tag{11}$$

式 (2) は

$$Ax' = b - Ax_0 \equiv b' \;. \tag{12}$$

したがって、問題は拘束条件 (11) の下に、式 (12) を解くことに帰着する。C の行ベクトルを $\{c^{(1)}, c^{(2)}, \ldots, c^{(\ell)}\}$ とすると、$(c^{(i)} \cdot x') = 0 \; (1 \le i \le \ell)$, すなわち、x' は $\{c^{(1)}, \ldots, c^{(\ell)}\}$ で張られる

部分空間 $\langle c^{(1)}, \ldots, c^{(\ell)} \rangle$ の直交補空間に属する。x' は C の核 $kerC$ に属する、すなわち $kerC$ の中で式 (12) を解くことになる。そのために、A の行ベクトル $\{a^{(1)}, \ldots, a^{(n)}\}$ を $\{c^{(1)}, \ldots, c^{(\ell)}\}$ に対して直交化する。その前にまず $\langle c^{(1)}, \ldots, c^{(\ell)} \rangle$ の正規直交基底を求めておく (例えば、Gram-Schmidt、あるいは SVD)。それらを $\langle \bar{c}^{(1)}, \ldots, \bar{c}^{(\ell)} \rangle$ とする。$\{c^{(1)}, \ldots, c^{(\ell)}\}$ が 1 次独立とは限らないので基底の数は ℓ 個とは限らないが、ここでは簡単のために基底の数は ℓ 個としておく。次に $\{a^{(1)}, \ldots, a^{(n)}\}$ を $\langle \bar{c}^{(1)}, \ldots, \bar{c}^{(\ell)} \rangle$ に関して直交化する：

$$\bar{a}^{(i)} = a^{(i)} - \sum_{j=1}^{\ell} \frac{(a^{(i)} \cdot \bar{c}^{(j)})}{\|\bar{c}^{(j)}\|^2} \bar{c}^{(j)} , \quad (i = 1, \ldots, n) . \tag{13}$$

これから、

$$(\bar{a}^{(i)} \cdot \bar{c}^{(j)}) = 0 , \quad (j = 1, \ldots, \ell), (i = 1, \ldots, n)$$

すなわち

$$\bar{a}^{(i)} \in kerC , \quad (i = 1, \ldots, n) \tag{14}$$

が成り立つ。また式 (13) から、$x' \in kerC$ に対して

$$(\bar{a}^{(i)} \cdot x') = (a^{(i)} \cdot x')\ , \quad (i = 1, \ldots, n) \quad \text{for } x' \in kerC$$

即ち $x' \in kerC$ に対して、$\bar{A}x' = Ax'$ が成り立つ。
また、方程式 $\bar{A}x' = b'$ の SVD 解は $kerC$ に属する。
（証）$\bar{A}$ の特異値分解を $\bar{A} = UWV^t$ とし、SVD 解を x' とすると、

$$x' = VW^+U^tb' = w_1^+(u^{(1)} \cdot b')v_1 + \cdots + w_m^+(u^{(m)} \cdot b')v_m$$

とかける。$w_i \neq 0$ に対応する v_i について

$$v_i \in (ker\bar{A})^\perp = \langle \bar{a}^{(1)}, \ldots, \bar{a}^{(n)} \rangle^{\perp\perp} = \langle \bar{a}^{(1)}, \ldots, \bar{a}^{(n)} \rangle \qquad (15)$$

式 (14)(15) から、$(v_i \cdot \bar{c}^{(j)}) = 0$、即ち $v_i \in kerC$。故に $x' \in kerC$
(証終)
以上から、$\bar{A}x' = b'$ の SVD 解 x' は束縛条件 (11) を満たし、かつ、
$\bar{A}x' = Ax'$ である。このとき SVD 解の性質により、
$\|Ax' - b'\| = \min$ あるいは、$Ax' - b' = 0$ の場合 $\|x'\| = \min$ となる。

3.1　例

問題

$$\begin{cases} x_1 - x_2 + x_3 = 1 \\ -x_1 + x_2 - x_3 = 0 \end{cases}$$

を拘束条件

$$x_1 + x_2 + x_3 = 0$$

の下に解け。

答　Schmidt 直交化法で解いてみる。

拘束条件を満たす解のうちの一つを $x_0 = \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}^t$ とすると、

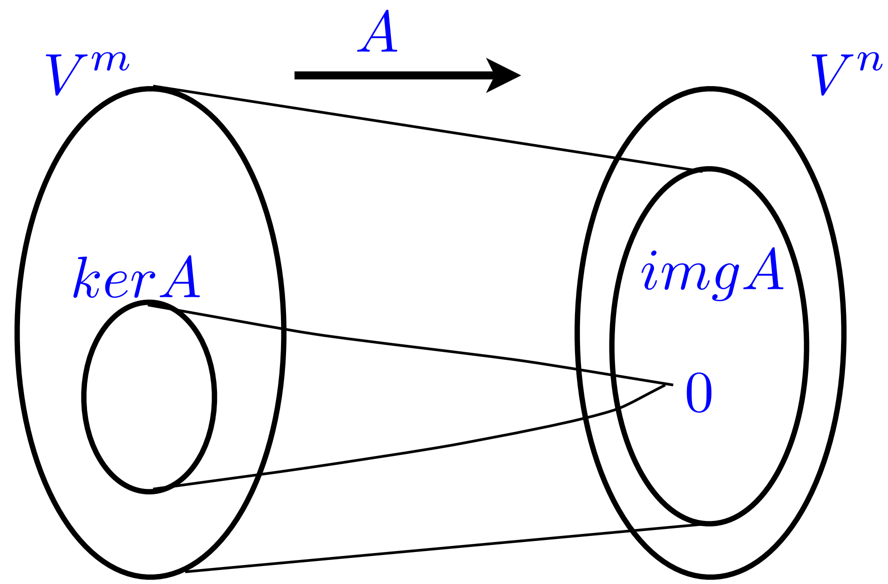
$b' = b - Ax_0 = b = (1)$、$x = x0 + x' = x'$。$A = \begin{pmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \end{pmatrix}$

を $C = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}$ について直交化すると、

$$\bar{A} = \begin{pmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \end{pmatrix} - \begin{pmatrix} 1/3 \\ -1/3 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{2}{3} & -\frac{4}{3} & \frac{2}{3} \\ -\frac{2}{3} & \frac{4}{3} & -\frac{2}{3} \end{pmatrix}$$

$\bar{A}x' = b'$ の SVD 解より、$x = x' = \begin{pmatrix} 0.125 & -0.25 & 0.125 \end{pmatrix}^t$ を得る。

補遺



$imgA := \{y \in V^n | \exists x \in V^m, Ax = y\}$

A　復習

大きさ (n, m) の行列 A と m 次元ベクトル $x(= (x_1, \cdots, x_m))$ を考える。A はベクトル空間 V^m からベクトル空間 V^n への 1 次写像を表す。A の列ベクトルを $\{a_1, \cdots, a_m\}$、A の行ベクトルを $\{a^{(1)}, \cdots, a^{(n)}\}$ と書く。

$$Ax = x_1 a_1 + \cdots + x_m a_m$$

と書けることから、A の像 $imgA$ はベクトルの集合 $a_i(i = 1, \cdots, m)$ によって張られる、すなわち、$a_i(i = 1, \cdots, m)$ の適当な線型結合をとれば $imgA$ の基底とすることができる。ベクトルの集合 $a_i(i = 1, \cdots, m)$ によって張られる部分空間を $\langle a_1, \cdots, a_m \rangle$ と書くことにすれば、

$$imgA = \langle a_1, \cdots, a_m \rangle .$$

また、Ax は

$$Ax = \begin{pmatrix} a^{(1)} \cdot x \\ \vdots \\ a^{(n)} \cdot x \end{pmatrix}$$

$ker A := \{x \in V^m | Ax = 0\}$

AのNull spaceともいう

とも書ける。これから、$Ax = 0$ を満たす x は $(a^{(1)} \cdot x) = \cdots = (a^{(n)} \cdot x) = 0$、すなわち $x \in \langle a^{(1)}, \cdots, a^{(n)} \rangle^{\perp}$。ここで $\perp$ は直交補空間を表わす（すなわちベクトル空間 V とその部分空間 W について、$W^{\perp} = \{x \in V | \forall y \in W, (x \cdot y) = 0\}$、また V は W と $W^{\perp}$ の直和となる：$V = W \oplus W^{\perp}$)。したがって、A の核 $ker A$ は

$$ker A = \langle a^{(1)}, \cdots, a^{(n)} \rangle^{\perp} .$$

B　特異値分解

定理 B.1 大きさ $(n,m),(n\geq m)$ の任意の行列 M は次のような形に分解できる。

$$M = UWV^t \tag{B1}$$

ここで U,W,V は次の型の行列である：$U(n,m),W(m,m),V(m,m)$。W は対角行列でその成分は正または零。U,V は列に関して正規直交（V は行に関しても正規直交）である。

$$u_i \cdot u_j = \delta_{ij} \qquad or \qquad U^tU = I \tag{B2}$$

$$v_i \cdot v_j = v^{(i)} \cdot v^{(j)} = \delta_{ij} \qquad or \qquad V^tV = VV^t = I \tag{B3}$$

W の対角成分を特異値、分解 (B1) を特異値分解（singular value decomposition; SVD）という。分解は特異値が相異なる場合、列の

置換を除いて一意的である。

$$
\begin{aligned}
Mx &= UW \begin{pmatrix} (v_1 \cdot x) \\ \vdots \\ (v_m \cdot x) \end{pmatrix} = U \begin{pmatrix} w_1(v_1 \cdot x) \\ \vdots \\ w_m(v_m \cdot x) \end{pmatrix} \\
&= w_1(v_1 \cdot x)u_1 + \cdots + w_m(v_m \cdot x)u_m
\end{aligned} \tag{B4}
$$

式 (B4) から次のことが分かる。

- $\{u_i | w_i \neq 0\}$ は M の像 $imgM$ の正規直交基底をなす。
- $\{v_i | w_i = 0\}$ は M の核 $kerM$ の正規直交基底をなす。
  $\because)$ $\{v_1, \cdots, v_m\}$ は V^m の正規直交基底であるから、$Mx = 0$ を満たす任意の x は

$$x = c_1 v_1 + \cdots + c_m v_m$$

とかける。$\{u_1, \cdots, u_m\}$ は 1 次独立であるから、式 (B4) より、

$$Mx = 0 \Longleftrightarrow w_i(v_i \cdot x) = 0 \ \ (i = 1, m)$$

であるが、もしある $j(1 \leq i \leq m)$ について $w_j \neq 0$ ならば $(v_j \cdot x) = 0$、即ち $c_j = 0$ である。したがって、x は $w_i = 0$ に対応する v_i の 1 次結合で書ける。

- $dim(imgM) + dim(kerM) = m$

系 B.2 $n < m$ の場合も同様に特異値分解できる。ただし、特異値は $w_i = 0, (i > n)$ であり、U の列ベクトルは $u_i = 0, (i > n)$ となる。U の列に関する正規直交性は $i \leq n$ に関してのみ成り立つ。

B.1 特異値分解による解の性質

x に関する方程式

$$Ax = b \tag{B5}$$

を A の特異値分解を利用して解く事を考える。A の特異値分解を

$$A = UWV^t \tag{B6}$$

とすれば、形式的に

$$\begin{aligned} x &= VW^{-1}U^t b \\ &= w_1^{-1}(u_1 \cdot b)v_1 + \cdots + w_m^{-1}(u_m \cdot b)v_m \ . \end{aligned} \tag{B7}$$

特異値が零の場合、式 (B7) の右辺は $ker A$ の方向に発散する。式 (B5) の x に $ker A$ のベクトルを加えても式 (B5) は変化しないが、その解は $ker A$ に近づく（大きさは無限大)。これは行列 A が特異（ill-conditioned）であることを表わしており、零あるいは零に近い特異値が数値的不安定をもたらしていると考えることができる。数値的不安定を緩和するために次の **SVD 解**を採用すると便利である。

$$\begin{aligned} x &= VW^{+}U^t b \ , \\ &= w_1^{+}(u_1 \cdot b)v_1 + \cdots + w_m^{+}(u_m \cdot b)v_m \ , \end{aligned} \tag{B8}$$

$$
\begin{aligned}
W^+ &= diag(w_1^+, \ldots, w_m^+) \ , \\
w_i^+ &= \begin{cases} w_i^{-1} & \text{if} \quad w_i \neq 0 \\ 0 & \text{if} \quad w_i = 0 \ . \end{cases}
\end{aligned}
\tag{B9}
$$

SVD 解は次の性質を持つ。

- b が $imgA$ に属するとき、解 x は $kerA$ の分だけ不定であるが、SVD 解はそれらの中でも最小のノルムをもつ。
（証）まず SVD 解は解の一つであることを示す。SVD 解 x に対して、

$$
\begin{aligned}
Ax = U(WW^+)U^t b &= w_1 w_1^+ (u_1 \cdot b) u_1 + \cdots + w_m w_m^+ (u_m \cdot b) u_m \\
&= \sum_{i \in \Lambda} (u_i \cdot b) u_i \qquad (\Lambda = \{i \mid w_i \neq 0\}) \ .
\end{aligned}
$$

上式の右辺は $w_i \neq 0$ なる i についての和であるが、対応する u_i は $imgA$ の直交基底をなすので、$b \in imgA$ より右辺は b に等しい。任意の解は $x + x'$ とかける。ここで x は SVD 解、x' は

$kerA$ の任意のベクトルである。x' は $kerA$ の基底 $\{v_i|w_i = 0\}$ の 1 次結合でかける。一方 SVD 解は $\{v_i|w_i \neq 0\}$ の 1 次結合となっている。$\{v_i \mid i = 1, \ldots, m\}$ は正規直交であるから、$x + x'$ のノルムは $x' = 0$ のとき最小となる。(証終)

- b が $imgA$ に属さないとき、$Ax = b$ は解を持たないが、SVD 解は $\|Ax - b\|$ を最小にする。

  （証）$b \notin imgA$ より、b は $imgA$ の元と $(imgA)^{\perp}$ の元との和でかける。

$$b = \sum_{i \in \Lambda} (u_i \cdot b) u_i + b' \ , \quad b' \in (imgA)^{\perp} \ .$$

  この表現は一意的である。右辺第 1 項は Ax（x は SVD 解）に等しい。$\bar{x} = x + x'$（x' は任意のベクトル）とすると、

$$A\bar{x} - b = Ax - b + Ax' = Ax' - b'$$

  となる。$Ax' \in imgA$、$-b' \in (imgA)^{\perp}$ であるから、$\|A\bar{x} - b\|$ は $Ax' = 0$ のとき、特に $x' = 0$ のときに最小となる。(証終)

式 (B8)(B9) において特異値が厳密に零でなくとも、数値的安定性を増すために、ある値以下より小さい特異値を強制的に零として SVD 解とすることがしばしば行われる。これは新しい特異値

$$\tilde{w_i} \quad = \quad \begin{cases} w_i & \text{if} \quad w_i \geq \epsilon \\ 0 & \text{if} \quad w_i < \epsilon \end{cases}$$

をもった新たな行列 $\tilde{A} = U\tilde{W}V^t$ の SVD 解を求めていることにほかならない。$\tilde{A}$ は元の行列 A の特異性（singularity）をある程度まで排除した、A の近似と考えることができる。$\tilde{A}$ はある意味で A の「最良」の近似である。すなわち、ある一定の階数 r（非零の特異値の数）をもつ行列の中では $\tilde{A}$ がフロベニウスノルム（Frobenius norm=行列要素の二乗和の平方根）の意味で最も A に近い。

http://www.youtube.com/watch?v=JEYLfIVvR9I&feature=player_profilepage